Chapter 3
Good Experimental Methodologies for Autonomous Robotics: From Theory to Practice

Francesco Amigoni, Viola Schiaffonati and Mario Verdicchio

Abstract A lively discussion on good experimental methodologies has recently taken place in the autonomous robotics community. Workshops have been organized, special issues published, and projects funded after recognizing that experimentation in autonomous robotics has not yet reached a level of maturity comparable with other fields of engineering and science. Within this discussion, we are not aware of any systematic survey on how experiments are conducted in autonomous robotics papers published in major journals and conferences. In this chapter, we provide an initial contribution to fill this gap by analyzing the experimental trends that emerge from the autonomous robotics papers presented over the last 11 years at the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), which constitute a privileged sample to get a picture on the evolution of experimental activity in the area. We conduct our analysis in the light of some principles that have been proposed for the development of good experimental methodologies in autonomous robotics.

Keywords Autonomous robots · Experimental methodologies

F. Amigoni (✉) · V. Schiaffonati
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy
e-mail: francesco.amigoni@polimi.it

V. Schiaffonati
e-mail: viola.schiaffonati@polimi.it

M. Verdicchio
Dipartimento di Ingegneria,
Università degli Studi di Bergamo, Bergamo, Italy
e-mail: mario.verdicchio@unibg.it

3.1 The Gap Between Theory and Practice

A lively discussion on good experimental methodologies has recently taken place in the autonomous robotics community. Workshops have been organized at major conferences [1], special issues published [2], and projects funded [3–6], which reflect the recognition that experimentation in autonomous robotics has not yet reached a level of maturity comparable with other fields of engineering and science. The interest in experimental methodologies is also motivated by the idea that they could reduce the gap between industrial applications and those applications, like service robotics, that require a significantly higher level of autonomy. If industrial robotics has some standards to evaluate systems employed in factories, the role of standardized experimental evaluation in autonomous robotics has not yet been universally recognized [7]. Within the context of this discussion, we are not aware of any systematic survey of the experimental activities presented in autonomous robotics papers published in major journals and conferences.

In this chapter, we aim at contributing to fill this gap by considering the autonomous robotics papers presented at the International Conference on Autonomous Agents and Multiagent Systems (AAMAS) over the last 11 years (from the inaugural edition in 2002 to the one in 2012). The selection of this source is motivated by a number of factors. Firstly, AAMAS is obviously a showcase of autonomous robotics research and features papers by many authors who regularly publish also at International Conference on Robotics and Automation (ICRA), International Conference on Intelligent Robots and Systems (IROS), and other robotic venues. However, differently from ICRA and IROS, autonomous robotics papers are well identified: they are usually presented in dedicated sessions. Moreover, considering papers from AAMAS is interesting to account for the point of view on experiments of researchers at the intersection between robotics and autonomous agents, which is a privileged perspective for observing autonomous robotics.

In this chapter, we analyze the trends that emerge from the experimental activities reported in 95 robotics papers presented at AAMAS in the light of some principles that have been proposed for the development of good experimental methodologies in autonomous robotics [8]. Although we are not claiming that they are the only principles that should be adopted in defining experimental methodologies for autonomous robotics, we deem that they are at the very foundation of experimental activity as traditionally intended and, hence, cannot be ignored. These principles are summarized below.

Comparison. The meaning of comparison in science is twofold: in a wider scope in the literature, it means to know what has already been done within the field, to avoid the repetition of uninteresting experiments and to get hints on promising issues to tackle; in a narrower scope of a specific kind of experiment, it refers to the possibility for researchers to accurately compare new results with old ones.

Reproducibility and repeatability. These features are related to the very general idea that scientific results should undergo to the most severe criticisms in order to be confirmed: reproducibility is the possibility for independent scientists to verify

the results of a given experiment by repeating it with the same initial conditions, instruments, and techniques, whereas repeatability is the property of an experiment that yields the same outcome from a number of trials, performed at different times and in different places.

Justification and explanation. This principle deals with well-justified conclusions based on the information collected during an experiment: it is not sufficient to collect as many precise data as possible, but it is also necessary to look for an explanation, that is, all the experimental data should be interpreted in order to derive the correct implications that lead to the conclusions.

Starting from the above principles, we identify a number of issues on how experimental activities are conducted. These issues include whether the papers present experiments, whether the experiments are performed in simulation or with real robots, whether the data or the code is available, whether the proposed system is compared with alternative systems, and so on. To the best of our knowledge, no other work has provided such an extensive survey over experimental activities in autonomous robotics. Our aim is not to present new methodologies, but to critically discuss some emerging trends that could help shed some light on the future of experiments in this field.

This chapter is organized as follows. The next section introduces the methodology we have used to select and analyze the papers we surveyed. Section 3.3 reports on the results of the analysis, while Sect. 3.4 discusses these results in the light of the experimental principles listed above. Finally, Sect. 3.5 concludes the paper.

3.2 Methodology

In this section we present the criteria we adopted in the process of selecting and analyzing the papers.

From the span of AAMAS conferences, from 2002 to 2012, we have considered all the papers that have been classified into robotics-related sessions. In this way, we avoided any arbitrary decisions in the selection of the papers. The titles of the sessions and the corresponding number of papers are reported in Table 3.1. In 2004 and 2007 no specific sessions on robotics were included in the proceedings. We ignored a 2007 session titled "Embodied agents and architectures", because it contains mostly papers related to synthetic and emotional agents. Overall, 95 papers have been analyzed. We explicitly note that these papers can be univocally identified in the AAMAS proceedings[1] starting from the information contained in Table 3.1.

A preliminary analysis of posters and short papers showed that there is no significant difference in the quality of the reported experimental activities with respect to full papers, although the number of experiments presented is obviously limited due to space constraints. Thus, since we are interested in analyzing the quality and not the quantity of experiments, posters and short papers have been included in our survey. Table 3.2 shows a classification of the 95 papers with respect to some of the keywords listed in the AAMAS2013 call for papers.

---

[1] Available at http://www.ifaamas.org/.

Table 3.1 Robotics papers at AAMAS

Year	Session	Number of papers
2002	Mobile embodied agents	9
	Robot architectures	4
2003	Robotics	4
2005	Robotics	5
	Robotics posters	3
2006	Robotics	10
2008	Multi-robotics full	8
	Multi-robotics short	11
2009	Multi-robotics	6
2010	Robotics I	6
	Robotics II	6
2011	Robotics	5
	Robotics and learning	4
2012	Robotics I	5
	Robotics II	5
	Robotics III	4

Table 3.2 Topics of robotics papers at AAMAS

Topic	Number of papers
Integrated perception, cognition, and action	12
Machine learning for robotics	10
Mapping, localization and exploration	6
Robot planning	32
Robot teams, multi-robot systems	5
Robot coordination	21
Robotic agent languages and middleware for robot systems	8
E-learning	1

The 95 papers we considered usually present techniques and methods that are applied to specific applications and problem settings, as shown in Table 3.3. Beyond generic navigation tasks that are fundamental for autonomous mobile robots, recurring applications include robotic soccer, surveillance and security (e.g., target recognition and tracking in military settings), and group behaviors like flocking. Other applications, aggregated in the final row of Table 3.3, include damage detection, self-assembling, robot dancing, and visitor companion (and e-learning). Note that two papers of 2005 and 2008, respectively, address two applications each.

We have considered a number of issues on how experimental activities are conducted to assess the properties of the proposed autonomous robotic systems. These issues are general enough to be applied to the wide range of topics the surveyed papers deal with. The list below is not definitive nor exhaustive, but it reflects our understanding of the issues characterizing experimental activities at the intersection between robotics and autonomous agents.

Table 3.3 Applications of robotics papers at AAMAS

Topic	Number of papers	2002	2003	2005	2006	2008	2009	2010	2011	2012
Environmental monitoring	2					1			1	
Exploration	7	1		1		1	1		1	2
Flocking and formation	11					5	1	2		3
Foraging	3			1		1				1
Generic navigation (e.g., coverage)	24	2		1	3	4	2	5	4	3
Locomotion (e.g., walking)	3	1	1		1					
Manipulation	2					1			1	
Object transportation	4	2	1	1						
Rescue	3					1		1	1	
Robotic soccer	12	2			3	4			1	2
Surveillance (e.g., patrolling and pursuit)	10	1		2		2	1	3		1
Target recognition and tracking	9	2	1	2	3					1
Other applications	7	2	1	1			1	1		1

Experiments. Does the paper present any experimental activity? By *experiments* we mean activities that require the implementation and the run of a computing system. In this sense, simple illustrative examples that are "drawn on paper" are not considered experiments for the purposes of this work. Moreover, we considered only what is reported in the paper: a paper does not qualify as a work that presents experiments if the only experimental activities discussed are illustrated in another work or are only claimed without explicit results.

Simulation or real robots. Among the papers describing experiments according to the above-mentioned criterion, we distinguish between simulations and activities with real robots. We further make a distinction between *standard* and *custom platforms*, that is, commercial and publicly available platforms, and the ones that are usually not made available outside the lab that developed them, respectively.

Data/code availability. Does the paper make *data* and *code* regarding the presented experiments available (e.g., downloadable from a website)?

Comparison. By *comparison* with other systems, we mean whether the proposed system is experimentally compared with others that perform the same function. We further distinguish between systems compared with *variants* developed by the same authors or with baseline systems, and systems compared with *alternatives* developed by other authors. An example of the former case is a learning system evaluated with or without a deadline for learning, while the latter case includes comparing the performance of the proposed learning system with that of Q-learning, for example. Please

note that we consider comparisons with baseline systems, like random methods in case of task assignment, as comparisons with variants developed by the same authors.

Measures. We check what is *measured* in the experiments, with a classification into two broad categories: *effectiveness* (or *functional*) measures, i.e., measures of what a system is supposed to do, and *efficiency* (or *non-functional*) measures, i.e., measures of the resources the system requires to obtain such performance. Examples of functional measures include the rewards obtained by robots in a reinforcement learning setting, the probability of detecting an intruder in a patrolling application, and success rate in reaching some desired configuration for swarm robotic systems. Examples of measures that we consider non-functional are the communication overhead associated to a given performance, the rate of convergence for a learning algorithm, and all measures of time and space complexity.

Settings. We consider the *settings* in which experiments are performed, and make a distinction between experiments performed in a single setting (e.g., one environment for a navigation algorithm or one task for a learning algorithm) and in different settings (e.g., multiple environments or multiple tasks).

Statistical analysis. Finally, we focus on the way in which the experimental activities reported in papers deal with *uncertainty* and *randomness* that affect all robotic systems operating in dynamic environments: are papers just presenting averages (or medians) or are they also presenting a more robust statistical analysis (e.g., based on ANOVA)? We consider papers showing at least standard deviation as presenting (a very simple, indeed) statistical analysis.

We will discuss in Sect. 3.4 how these issues are related to the experimental principles mentioned in Sect. 3.1.

As in any survey, our selection of issues is subjective and amendable. However, the issues we consider are so general that they can be applied to all the topics in Table 3.2. Moreover, our analysis of the papers presents some inescapable limitations. For example, sometimes multiple systems are presented in the same paper, that could be classified in different ways: some of them are compared with alternatives, whereas some others are not. In such cases, we chose to go for the best available option (e.g., we considered the whole paper as comparing the proposed system with alternatives if it is done for at least one of the proposed systems). Notwithstanding its limits, we believe that the analysis reported in the next section contributes to form an initial picture of the experimental trends in autonomous robotics.

3.3 Analysis

In this section, we present the results of our analysis of the AAMAS robotics papers and try to identify some emerging trends in the experimental activities of autonomous robotics. The results of our analysis are presented in graphs, showing the fraction of papers addressing the issues illustrated in the previous section.

Let us start from Fig. 3.1, which shows that the majority of papers present experiments. This should not come as a surprise, as experimentation is the main way in

papers reporting experiments
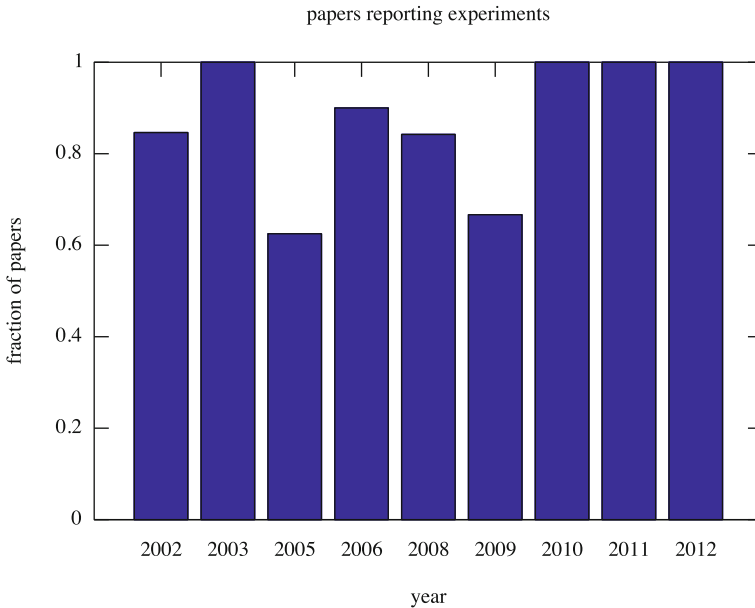
Fig. 3.1 Fraction of AAMAS robotics papers that present experiments

which robotic systems are evaluated and assessed. In some years of the surveyed time span, all papers present experimental activities, whereas in some others such fraction is lower. For example, in 2005, 3 out of 8 papers do not present experiments; 2 (short, 2-page) papers claim that experiments have been performed, but do not describe them, while another (full) paper is theoretical and presents only simple examples of application of the proposed algorithm for pursuit evasion. It is interesting to notice that all papers in the last 3 years (2010, 2011, and 2012) present experiments.

Obviously, the way in which experiments are intended strongly varies (also according to the topic of the paper): they range from simple qualitative descriptions of the behavior of the implemented systems to more sophisticated evaluation activities, involving different alternative approaches and possibly supported by statistical analysis. Let us then try to shed some light on how experiments are performed.

Figure 3.2 shows the fraction of papers with experiments that use simulated and real robots. Note that a paper may present experiments with both simulated and real robots. In such a case, the paper contributes to both fractions, which is why the bars relative to a year may sum up to more than 1. Simulation dominates over real robots, which can be explained by the lower costs and the relatively easier operational aspects of simulation. However, it is interesting to notice that the fraction of papers presenting experiments with real robots is somehow constant over the years. This could be related to the fact that papers addressing some common topics, like target tracking, are more frequently presenting experiments with real robots. A common

simulated (blue) or real (red) robots
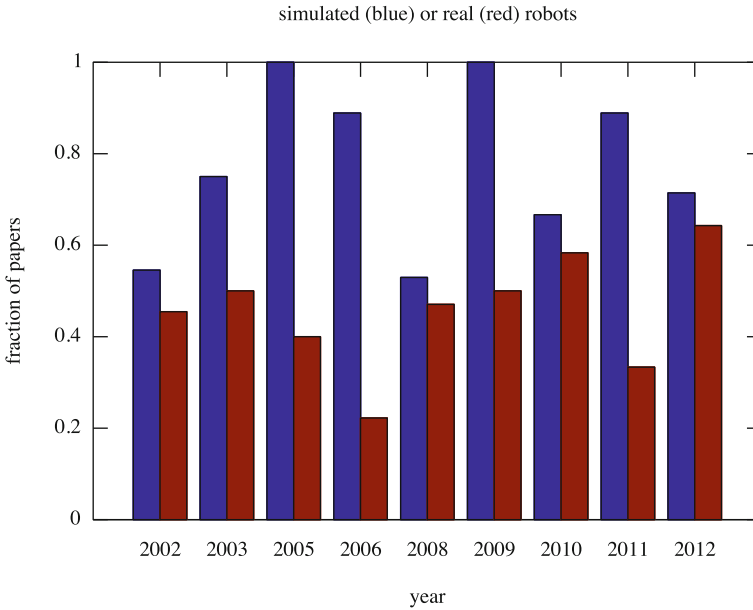
Fig. 3.2 Fraction of papers that use simulated and real robots in experiments

**Table 3.4** Standard simulators used in robotics papers at AAMAS

Simulator	Number of papers	2002	2003	2005	2006	2008	2009	2010	2011	2012
(Player/) Stage or Gazebo	4				1				1	2
Robocup simulators (e.g., USARSim)	6				1	1		1	2	1
Cyberbotics' webots	6	2			1	1		1	1	

situation is that in which extensive experiments are performed in simulation and simpler demonstrations are performed with real robots.

Figure 3.3 shows the analysis of the simulation tools, with a particular focus on standard simulators as opposed to custom ones. The use of standard simulators seems to be increasing over the years, which could be related to the fact that more and more reliable simulation platforms have recently become available. Looking at the standard simulators used in the last years, it emerges that most of them are used in competitions like RoboCup (e.g., USARSim [9]). Another standard simulator that has been used since 2002 and is still employed is Cyberbotics' Webots [10]. The standard simulators that have been used at least by 2 papers are reported in Table 3.4.

The trend of using standard platforms is even stronger in the case of experiments with real robots: Fig. 3.4 shows that there is an evident tendency in adopting standard

standard (blue) or custom (red) simulators
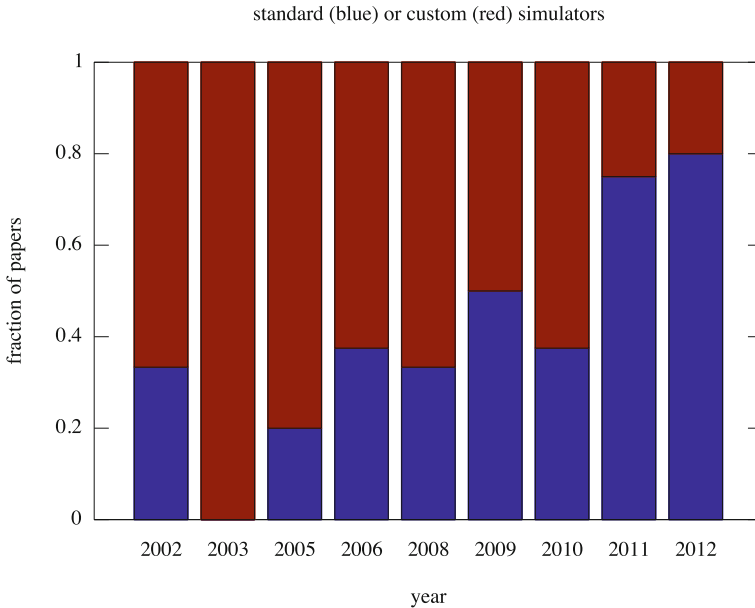
Fig. 3.3 Fraction of papers that employ standard and custom simulators

platforms in this type of research. A possible explanation is that standard robotic systems are usually easier to set up, more reliable, and, in some cases, cheaper in terms of resources and time than custom systems. Among the standard platforms that are employed more often, there are MobileRobots' Pioneer robots [11] and Sony's Aibo robots [12]. The standard robotic platforms that have been used at least by two papers are reported in Table 3.5.

Figure 3.5 shows the fraction of the papers that, whether with simulations or with real robots, compare their proposal with other systems. This analysis is interesting because it shows that, in a number of cases, robotic systems presented at AAMAS are only empirically shown to work, but such experiments do not include a comparison with other systems, which in some sense weakens their assessment.

Among those papers that provide experimental comparisons, some take only simple variants of the proposed systems into account, whereas other papers consider fully alternative systems, typically developed by other researchers. Figure 3.6 shows these two trends: it is evident that there is an increasing tendency toward a more sophisticated notion of comparison.

Figure 3.7 shows the fraction of papers (among those that present experiments) presenting functional (effectiveness-related) and non-functional (efficiency-related) experimental measures. If functional measures are more frequently presented (not surprisingly, since the goal of experimenting is to show that a proposed system works), the illustration of non-functional measures has reached a good diffusion in the last years. This allows for a more complete understanding and assessment of

standard (blue) or custom (red) real robotic systems
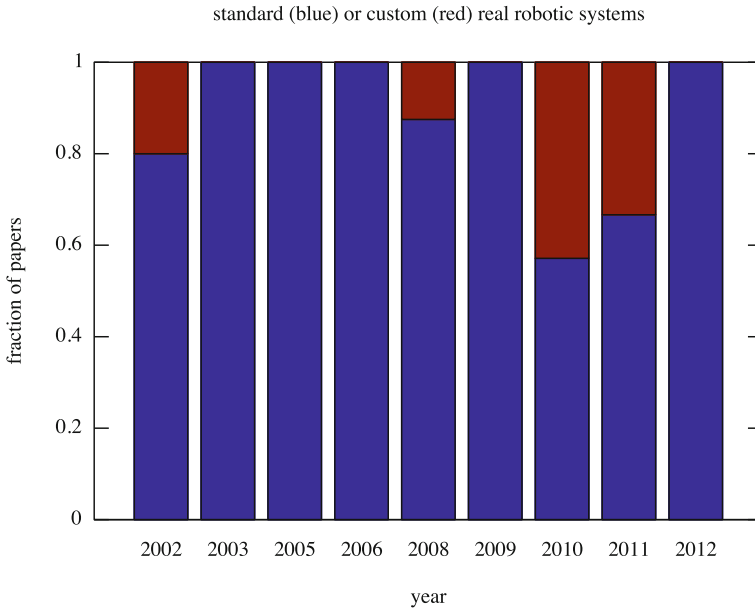
Fig. 3.4 Fraction of papers that employ standard and custom real robotic systems

**Table 3.5** Standard real robotic systems used in robotics papers at AAMAS

Robotic system	Number of papers	2002	2003	2005	2006	2008	2009	2010	2011	2012
Sony's AIBO	2					2				
KOVAN's kobot	2					1	1			
Lego's mindstorms	2		1		1					
Mobilerobots' pioneer family (I, 2-DX, 3-DX, P3-AT)	8	2	2		1		1	1		1
Friendly robotics' RV400	3					1			2	
Segway's RMP	2			1		1				
Albedaran's nao	3									3
EPFL's e-puck	2									2

robotic systems, by showing not only the success rate in the tasks a system has been designed for, but also the resources that are needed to achieve such results.

Figure 3.8 shows the papers (among those presenting experiments) that illustrate experiments in single or multiple settings. Naturally, an experimental activity conducted in multiple settings is expected to come to stronger conclusions than an experimental activity conducted in a single setting. However, sometimes the use of a single
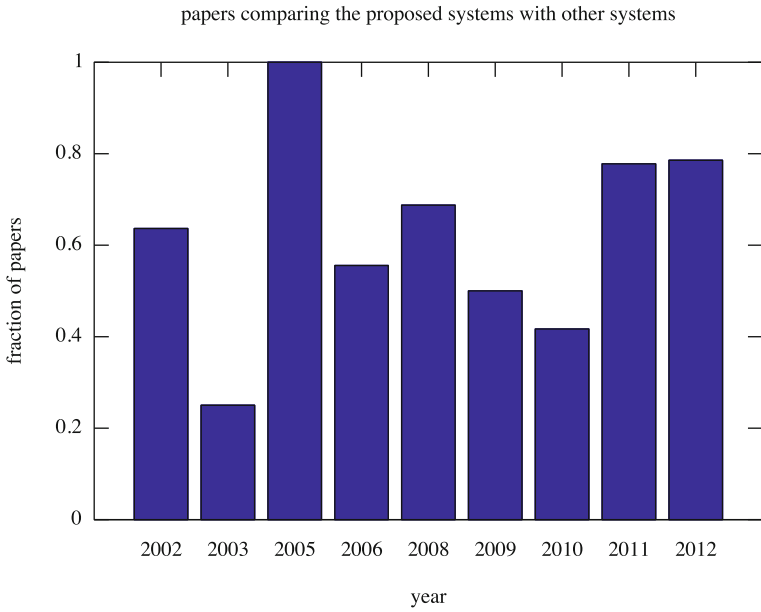
papers comparing the proposed systems with other systems

Fig. 3.5 Fraction of papers that compare the proposed systems with other systems

comparison with variants of the proposed system (blue) or with alternative systems (red)
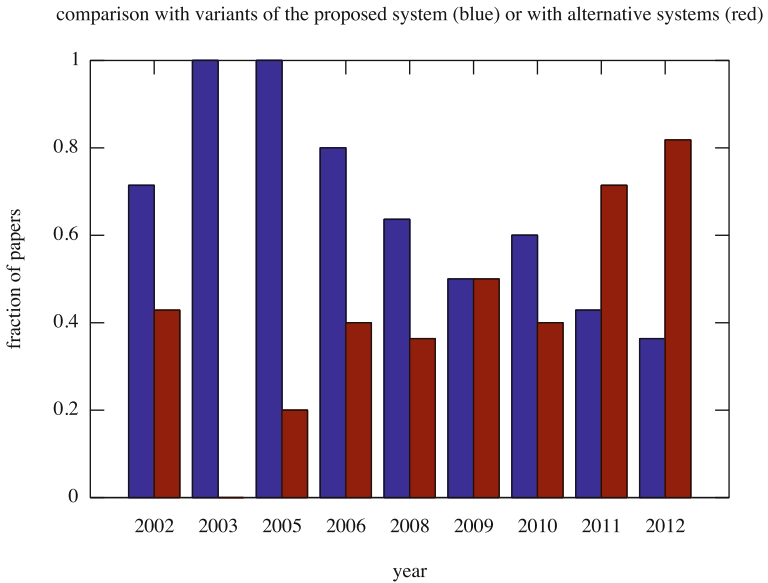
Fig. 3.6 Fraction of papers that compare the proposed systems with variants of the same systems or with alternative systems

functional (blue) or non-functional (red) measures
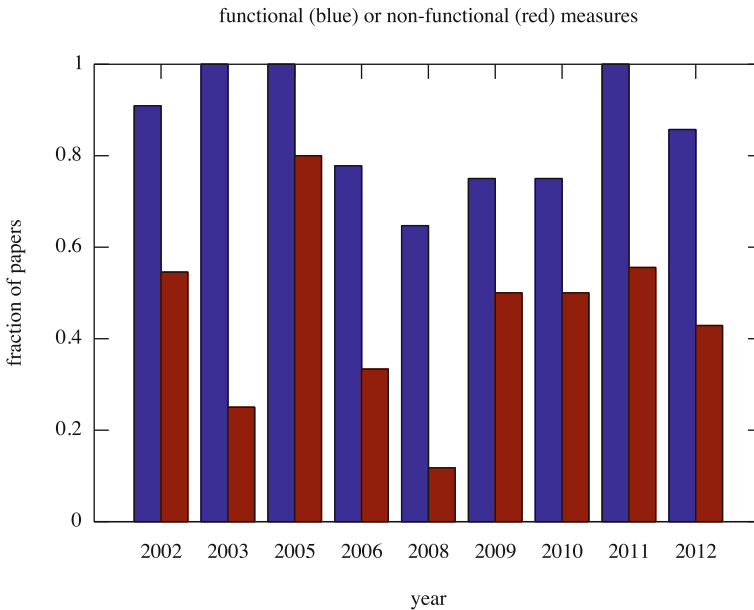
Fig. 3.7 Fraction of papers that present functional and non-functional experimental measures
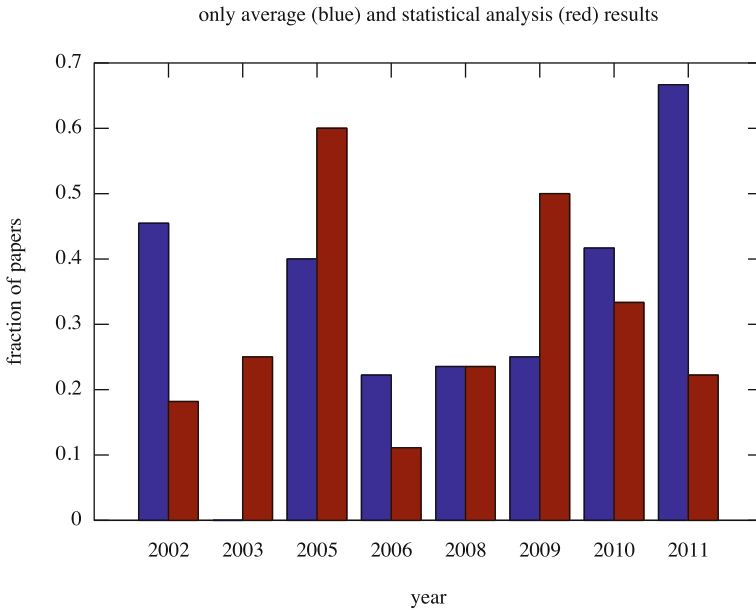
setting is inescapably "forced" by the application, like in the case of robotic soccer. The number of papers presenting experiments in multiple settings is indeed raising in the last years, which may be related to the increase in use of standard simulators (see Fig. 3.3) that make it easier to conduct experiments in different environments.

One of the typical characteristics of autonomous robotic systems is that they have to deal with uncertainty and randomness of unpredictable environments. Figure 3.9 shows the fraction of papers (among those presenting experiments) that present average-only and statistically analyzed experimental results. It emerges that randomness is not always dealt with in a statistically-sound way. For example, a paper proposing a novel navigation system presents an experimental activity involving a number of trials performed in real environments in which the robot was required to go from a random location to a target location. Reported results show only the average distance and speed for two different settings, making it difficult to precisely assess the significance of the findings and the difference between the two settings.

In our analysis, we found only few papers presenting links to available code and/or data. In particular, one paper in 2002, one in 2008, one in 2010, and two in 2012 provided links to additional material (like further experiments or videos of experiments), but the links appear to be broken. Other papers, however, i.e., one in 2003, two papers in 2008, three in 2010, and five in 2012, featured links to additional material which are working (as of April 2012).

single (blue) or multiple (red) settings
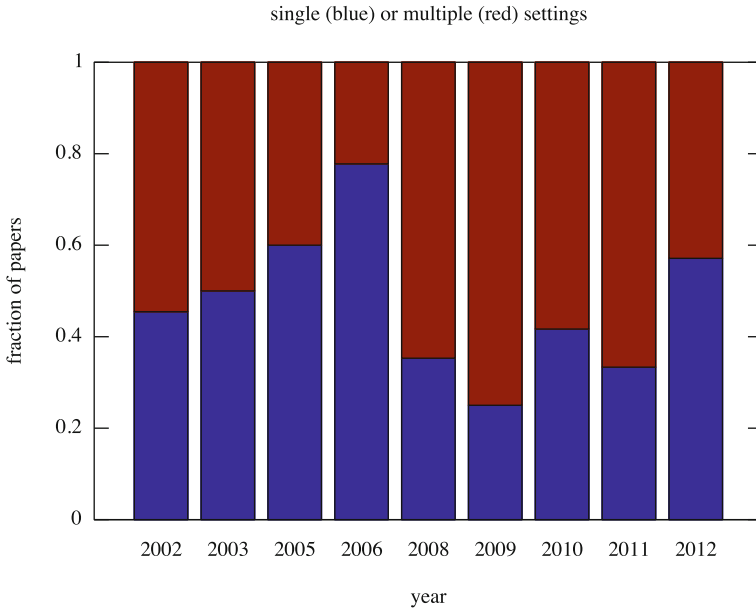
Fig. 3.8 Fraction of papers that present experiments in single or multiple settings

only average (blue) and statistical analysis (red) results

Fig. 3.9 Fraction of papers that present average-only and statistically-analyzed experimental results (note truncated scale on vertical axis)

Finally, we observed that very few papers (5 out of 81) report negative results. In these rare cases, examples and instances in which the proposed approach fails or for which it has a low success rate are presented.

3.4 Discussion

In the previous section we have presented the experimental trends emerging from our analysis of AAMAS robotic papers. Although not in all cases clear trends can be identified out of the figures, nevertheless we think that our analysis can point at some clues to keep an eye on, which, if confirmed by the papers published in the upcoming years, could provide us with the right perspective on the evolution of robotics as a more mature experimental discipline. The fact that trends do not always emerge in clear ways can be due to the great variability of the papers considered and also to the limited pool from which we have taken the papers: 11 years of research may not be enough to provide us with definite indications on the directions this field is going to take.

The issues we analyzed in the previous section are in accordance to the very idea of experiment as developed in science after the so called Scientific Revolution in the seventeenth century, and based on the principles listed in Sect. 3.1. An experiment is a *controlled experience*, namely a set of observations and actions, performed in a controlled context, to support a given hypothesis. Accordingly, better achievements can be gained if knowledge is at disposal and can be rigorously compared among scholars. This control is guaranteed by reproducibility and repeatability, as well as by the justification/explanation principle expressing the idea that experiments are not just a collection of observations and data, but require their explanation.

Here follows a discussion of some of the identified experimental trends in autonomous robotics, according to the general principles mentioned above. We note that following these principles means changing the perception of what an experiment is in the autonomous robotics community: from experiments intended just as ways to show, by means of few examples, that a given system works appropriately, to experiments where an important issue is to show how the system works and how its performance compares to alternative systems.

More experiments in recent years. All papers from 2010–2012 present experimental results. Two possible factors may have played a significant role. Firstly, the review process may have considered experiments as a requirement for a paper to pass the selection process for the AAMAS conference, so that, although research practices have not significantly changed in the community, experiment-oriented works tend to be chosen for publication over others. Secondly, there might be a shift in the community toward experimentation, so that, independently of the reviewing process, all in all, we have more experimentation in the field of robotics.

More simulation than real robots. Beyond the simpler explanation that simulations are cheaper than physical robots, this tendency might depend on the fact that simulators have become more and more sophisticated, enabling researchers to

model robots and environments with more detail and precision. The increased reliability of the simulation results may have triggered a change in the attitude of robotics researchers toward simulators, not seen as "fake robotics" anymore, but as first-class citizens of the field.

Increased use of standard platforms. Adopting a well-documented, existing tool is often a quicker and easier way to obtain results than starting from scratch. The above-mentioned increased reliability of simulators, and a similar phenomenon in the context of physical robots may have boosted the adoption of standard systems throughout the research community. This trend is also supported by efforts like ROS [13], which aims at providing a common platform for the control modules of the robots. The use of standard simulators and robots has surely a positive effect on the reproducibility and repeatability of experiments, so that they can be performed under different conditions and by several researchers, with the possibility to strengthen previously obtained results, or to put them under discussion.

Weakness of experimental comparison of systems. The fraction of papers that feature an experimental comparison of the proposed system with other systems is not very large. This phenomenon may be due to the difficulty to determine a common ground for experimental comparisons. For instance, identifying a set of standard environments for benchmarking navigation systems is an hard problem, let aside the identification of measures for comparing the systems. However, although the number of papers performing comparisons is not large, we have witnessed an increase in the quality of comparison, including more frequent evaluation of alternative systems developed by other researchers. This could be again connected to the adoption of standard platforms. For example, using RoboCup soccer simulators, it is easy to compare different teams by having them play against each other. Should this tendency be confirmed in the future, it would surely have a positive impact on the scientific aspects of autonomous robotics research, enhancing the comparison principle.

More attention toward non-functional parameters. Industrial robotics has had a tremendous impact on the industry and the economy, whereas more cutting-edge proposals coming from artificial intelligence and autonomous robotics have only recently found their way into the world outside the labs (e.g., with entertainment robots, household robots, prosthetics). This means that more and more products of research in robotics must meet tighter requirements given by the services these artifacts are supposed to provide in everyday life, which may have triggered in researchers a bigger interest in the efficiency of their systems. Moreover, a better understanding of the systems strengthens the justification/explanation principle toward a more scientific approach to experiments.

Little attention to statistical analysis. This trend evidences still a distance from the principle of justification/explanation mentioned above. A rigorous statistical analysis nowadays is a necessary requirement to treat data in a meaningful way if we want to look for real explanations instead of mere collection of data. A similar situation has been faced some years ago in artificial intelligence [14] and brought to a widespread use of statistical analysis in many areas of the field.

Low availability of data and code. This trend, again, is not in accordance with the general idea of experiments, according to which data should be compared in the

most rigorous and extensive way. In the papers we have analyzed, very few present full (raw) experimental data or make code available to the reader. Although in some cases there are copyright and privacy issues that prevent distribution of data and code, this shortcoming obviously weakens the possibility to replicate experiments by other researchers and hence to extensively control experimental results in an independent way. A possible way to overcome this problem could be the creation of a repository for additional material of the papers. Moreover, it could be useful to stimulate the use of publicly available datasets, like Radish [15] and RAWSEEDS [4], in experiments. A step in this direction could also address the complaint (sometimes encountered in AAMAS papers) that there are no standard benchmarks for evaluating autonomous robotic systems.

3.5 Conclusions

Our efforts in analyzing the AAMAS robotics papers of the last 11 years allowed us to spot some clues that make us optimistic about the future development of the experimental methodologies in this discipline. The increase in the use of standardized platforms, both for simulations and for real robots, seems to be the most promising of all. The convergence of the endeavors of several researchers toward the same platforms can have a positive impact on the significance of the results of their work, in that, they can be throughly verified under different conditions by different testers thanks to the standard systems that provide an easily accessible common ground to all the members of the community. Moreover, as simulators become more and more reliable, the financial constraints that come with physical robotic systems can be bypassed without compromising the verisimilitude of the experiments, thus allowing also researchers from smaller labs to join forces in the robotics enterprise. However, some other aspects seem more problematic and suggest that things should be done differently, like the scarce use of statistical analysis and the difficulty in making data and code publicly available to ease the comparison of systems.

We discussed only some trends that emerge from our analysis. Researchers can easily identify several other patterns that it is worth discussing. Future work could address a more complete analysis of how experiments are conducted in autonomous robotics, both considering more issues (possibly specialized for particular sub-fields) and a larger sample of papers. In particular, it could be interesting to compare the trends emerged from our analysis of AAMAS robotics papers to those of other conferences, like ICRA. To this end, a criterion to select autonomous robotics papers from the large number of papers in the ICRA proceedings is needed.

References

1. Euron GEM Sig (2007) http://www.heronrobots.com/EuronGEMSig/. Accessed July 2013
2. Madhavan R, Scrapper C, Kleiner A (2009) Special issue on characterizing mobile robot localization and mapping. Auton Robot 27(4):309–481
3. BRICS (2009) Best practice in robotics. http://www.best-of-robotics.org/. Accessed July 2013
4. Rawseeds (2006) http://www.rawseeds.org/. Accessed July 2013
5. RoCKIn (2013) Robot competitions kick innovation in cognitive systems and robotics. http://rockinrobotchallenge.eu/. Accessed July 2013
6. RoSta (2007) Robot standards and reference architectures. http://www.robot-standards.eu/. Accessed July 2013
7. Bonsignorio F, Hallam J, del Pobil AP (2007) Good experimental methodology - GEM guidelines. http://www.heronrobots.com/EuronGEMSig/downloads/GemSigGuidelinesBeta.pdf. Accessed July 2013
8. Amigoni F, Reggiani M, Schiaffonati V (2009) An insightful comparison between experiments in mobile robotics and in science. Auton Robot 27(4):313–325
9. USARSim (2011) Unified system for automation and robot simulation. http://usarsim.sourceforge.net. Accessed July 2013
10. Webots (2011) http://www.cyberbotics.com/overview. Accessed July 2013
11. Adept Technology (2011) http://www.mobilerobots.com/. Accessed July 2013
12. Sony Aibo (2006) http://support.sony-europe.com/aibo/. Accessed July 2013
13. ROS (2011) Robot operating system. http://www.ros.org/. Accessed July 2013
14. Cohen P (1995) Empirical methods for artificial intelligence. The MIT Press, Cambridge
15. Howard A, Roy N (2003) The robotics data set repository (radish). http://radish.sourceforge.net/. Accessed July 2013